Unsupervised Tumour Segmentation in PET using Local and Global Intensity-Fitting Active Surface and Alpha Matting

Zeng, Ziming; Wang, Jue; Tiddeman, Bernie; Zwiggelaar, Reyer

Published in:
Computers in Biology and Medicine

DOI:
10.1016/j.compbiomed.2013.07.027

Publication date:
2013

Citation for published version (APA):
Zeng, Z., Wang, J., Tiddeman, B., & Zwiggelaar, R. (2013). Unsupervised Tumour Segmentation in PET using Local and Global Intensity-Fitting Active Surface and Alpha Matting. *Computers in Biology and Medicine*, *43*(10), 1530-1544. https://doi.org/10.1016/j.compbiomed.2013.07.027

Unsupervised Tumour Segmentation in PET using Local and Global Intensity-Fitting Active Surface and Alpha Matting

Ziming Zeng*[a,b], Jue Wang[c], Bernie Tiddeman[a], Reyer Zwiggelaar[a]

[a] *Department of Computer Science, Aberystwyth University, Aberystwyth, UK*
[b] *Faculty of Information and Control Engineering, Shenyang Jianzhu University, Shenyang, China*
[c] *Adobe Systems, Seattle, USA*

Abstract

This paper proposes an unsupervised tumour segmentation approach for PET data. The method computes the volumes of interests (VOIs) with subpixel precision by considering the limited image resolution and partial volume effects. First, an improved anisotropic diffusion filter is used to remove image noise. A hierarchical local and global intensity active surface modelling scheme is then applied to segment VOIs, followed by an alpha matting step to further refine the segmentation boundary. The proposed method is validated on real PET images of head-and-neck cancer patients with ground truth from human experts, as well as customly-designed phantom PET images with objective ground truth. Experimental results show that our method outperforms previous automatic approaches in terms of segmentation accuracy.

Keywords:
Positron Emission Tomography, Tumour Segmentation, Active Surface Modelling, Alpha Matting, Mutual Information.

1. Introduction

Positron emission tomography (PET) is one of the most advanced medical imaging technologies on the molecular level for human functional imaging. A

*Corresponding author. Tel.: +44 (0)7438841361; Fax:+44 (0)1970 628536; Email: zengziming1983@gmail.com.

common example in clinical use is imaging glucose metabolism with the PET tracer ^{18}F-Fluorodeoxyglucose (FDG). High up-take and retention of FDG by tumour cells can be used to define a volume of interest (VOI). Accurate segmentation of the VOI is critical for many tasks in radiation oncology such as diagnosis, staging, radiotherapy planning and treatment assessment. In addition to the primary tumour, metastatic VOIs such as lymph nodes in head-and-neck cancer also express high FDG uptake, thus their segmentation is also important for staging and treatment planning.

Segmentation of PET VOIs is usually performed manually, or by thresholding at standardized uptake values (SUVs) [1]. Manual segmentation is time consuming and the VOIs delineated by different medical professionals can vary widely [2]. On the other hand, the majority of fully- or semi-automatic segmentation methods are based on intensity thresholding [1, 3, 4, 5, 6, 7]. It is difficult to find a threshold value to segment the VOI accurately due to the low spatial resolution and the signal-to-noise characteristics of PET, and the absence of a meaningful scale (such as Hounsfield units (HU) in CT imaging) to associate voxel intensities with underlying molecular processes. Also, thresholding is not very robust, as a small change of the threshold value can significantly alter the segmentation result. Last but not least, thresholding does not consider the spatial characteristics of an image, making it sensitive to noise and intensity inhomogenity.

Recently, more advanced automatic PET segmentation methods have been proposed. Belhassen et al. [8] employed a fuzzy C-means algorithm to segment tumours in PET. Hsu et al. [9] utilized an edge-based segmentation model called Poisson Gradient Vector Flow (PGVF) to automatically find the abnormality outlines in PET images. Gribben et al. [10] used the Markov Random Field Expectation Maximisation (MAP-MRF EM) labelling technique to segment VOIs. Montgomery et al. [11] proposed a segmentation method based on a multiscale Markov Random Field Model (MRFM). These methods only consider intensity and spatial characteristics of PET images, but ignore the information of the whole volume. Also, these methods are hindered by poor spatial resolution and partial volume effects (PVE).

In this paper, to address the above limitations of previous approaches, we propose a novel VOI segmentation method which employs a hierarchical approach. Our method (Section 2) combines improved local and global intensity active surface modelling and alpha matting, and achieves sub-pixel segmentation accuracy. Experimental results (Section 3) show that this approach is robust with respect to noise and density inhomogenity in PET data,

and can achieve accurate segmentation results.

2. Overview of the Proposed Approach

The flowchart of the proposed segmentation system is shown in Fig. 1. First, as preprocessing, an improved anisotropic diffusion filter is used to automatically remove image noise. A hierarchical segmentation scheme is then employed, which considers the global and local volumetric information. In the segmentation scheme an improved 3D active surface modelling method is used to segment the VOIs in the PET volume. Finally, a trimap which contains a definite foreground, a definite background and an unknown region is automatically generated and an alpha matting technique is used to refine the segmentation results.

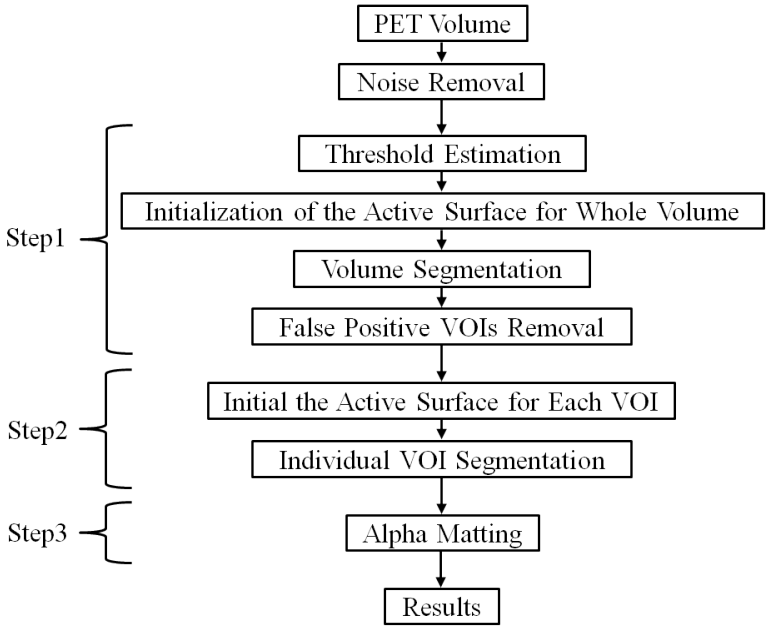

Figure 1: Flowchart of the proposed unsupervised PET image segmentation approach.

2.1. Noise Removal

To overcome the low signal-to-noise ratio of PET images, an anisotropic diffusion method is used to reduce noise in each PET slice. This method is able to effectively remove additive noise while retaining the image edges.

The traditional nonlinear diffusion filtering method proposed by Perona and Malik [12] can be described as:

$$\begin{cases} \frac{\partial I}{\partial t} = div[c\,(\mid \nabla I \mid) \cdot \nabla I], \\ I\,(t=0) = I_0, \end{cases} \tag{1}$$

where ∇ is the gradient operator, $div()$ is the divergence operator,$\mid \cdot \mid$ denotes the magnitude, $c\,(x)$ denotes the diffusion coefficient, and I_0 is the initial image. The diffusion coefficient which was improved in [13] is defined as:

$$c\,(x) = \frac{1}{ln[e + (x/\kappa)^2\,]}, \tag{2}$$

where κ is an edge magnitude parameter.

Under the control of $c\,(\mid \nabla I \mid)$, the model can achieve selective smoothing based on the gradient field of the original image. For edges with large gradient magnitudes, $c\,(\mid \nabla I \mid)$ is small, and the model will perform hardly any smoothing in order to keep the high frequency edge details. On the other hand, for flat areas where the gradient magnitudes are small, $c\,(\mid \nabla I \mid)$ will be large, and the model will perform a strong smoothing to remove noise. However, to allow this method to work well the iteration number needs to be specified by the user.

In our work, we have included an automatic stopping criteria to eliminate the requirement for user input. The flowchart of our new automatic noise removal scheme is show in Fig. 2. The method is based on estimating the difference in mutual information (MI) in successive steps. The iteration will stop when the difference ($\partial MI = \mid MI(k) - MI(k-1) \mid$) is smaller than an empirically determined threshold value (τ). In addition, the traditional anisotropic diffusion only considers 4 directional gradients (up, down, left and right). In our work we extend the anisotropic diffusion to eight directions for improved accuracy.

2.2. Improved 3D Active Surface Modelling

Active surface modelling is one of the most successful deformable model approaches for image segmentation, especially in the area of medical image analysis [14]. It can detect edges with sub-pixel accuracy and can easily be formulated into an energy minimization framework. In the 3D segmentation step, an intensity-based active surface modelling method [15, 16] which considers both local and global intensity statistics is used for segmenting the

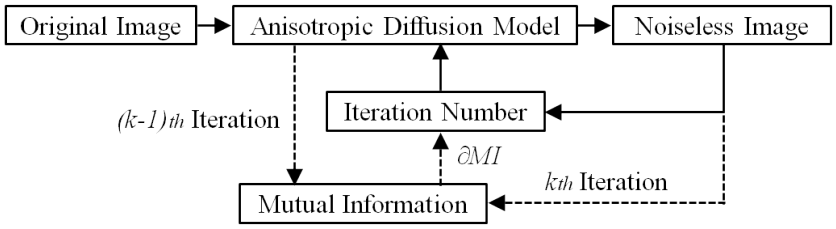

Figure 2: Flowchart of the automatic noise removal scheme.

whole volume. A globally-convex segmentation model [17] is employed to derive an energy function that can be minimized by convex optimization. Subsequently, the energy function is minimized along with the deformation of the surface using a Split Bregman technique [17]. As shown in [15], the energy function is formally defined as:

$$\mathcal{F}(\phi, c_1, c_2, f_1, f_2) = \mathcal{E}^{LGIF}(\phi, c_1, c_2, f_1, f_2) + \mu\mathcal{P}(\phi) + \upsilon\mathcal{L}(\phi), \tag{3}$$

$$\mathcal{E}^{LGIF}(\phi, c_1, c_2, f_1, f_2) = (1-\omega)\mathcal{E}^{LIF}(\phi, f_1, f_2) + \omega\mathcal{E}^{GIF}(\phi, c_1, c_2), \tag{4}$$

where ϕ denotes the tumour contour, $\mathcal{P}(\phi) = \int (1/2)(\mid \nabla\phi(x) \mid -1)^2 dx$ is the level set regularization term, $\mathcal{L}(\phi) = \int \mid \nabla H(\phi(x)) \mid dx$ is the smoothness term, ω is a positive constant in $[0, 1]$, $\mathcal{E}^{LIF}$ is the local energy function and $\mathcal{E}^{GIF}$ is the global energy function, which are defined as:

$$\mathcal{E}^{LIF}(\phi, f_1, f_2) = \sum_{i=1}^{2} \lambda_i \int [\int K_\sigma(x-y) \mid f_i(x) - I(y) \mid^2 M_i(\phi(y)) dy] dx, \tag{5}$$

$$\mathcal{E}^{GIF}(\phi, c_1, c_2) = \sum_{i=1}^{2} \lambda_i \int \mid I(x) - c_i \mid^2 M_i(\phi(x)) dx, \tag{6}$$

where $K_\sigma(u)$ is a 3D Gaussian kernel function. $K_\sigma(x-y)$ is a weight on point y with regard to the centroid point x of the Gaussian kernel. The local fitting energy $\mathcal{E}^{LIF}$ is determined by the value of σ. $M_1(\phi) = H(\phi), M_2(\phi) = 1 - H(\phi)$. The Heaviside function H is usually approximated by a smooth function $H_\epsilon(x) = (1/2)[1 + (2/\pi) arctan\,(x/\xi)]$. As the grey level of tumours can vary between volumes, we changed this function to attract the active surface to the boundary of the target object as:

$$H_\epsilon(x) = \frac{1}{2}\Big[1 + \frac{2}{\pi} arctan\left(\frac{x-T}{\xi}\right)\Big], \tag{7}$$

where T $(0 \leq T \leq 1)$ denotes a normalised tumour grey level threshold, which is the average of the maximum grey level value of some selected slices which will be introduced in Section 2.3.

According to the derivation by Wang et al. [15], the optimal functions $f_1(x)$, $f_2(x)$ and constants c_1, c_2 that minimize $\mathcal{F}(\phi, c_1, c_2, f_1, f_2)$ defined in Eq. 3 are:

$$f_i(x) = \frac{K_\sigma(x) * [M_i(\phi(x))I(x)]}{K_\sigma(x) * M_i(\phi(x))}, \;\; i = 1, 2 \tag{8}$$

$$c_i(x) = \frac{\int I(x) M_i(\phi(x)) dx}{\int M_i(\phi(x)) dx}, \;\; i = 1, 2. \tag{9}$$

The optimal fitting functions $f_1(x)$, $f_2(x)$ and the constants c_1, c_2 that minimize the energy function ϕ can be derived from:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\Big[\upsilon div\Big(\frac{\nabla \phi}{\mid \nabla \phi \mid}\Big) - (F_1 + F_2)\Big] + \mu\Big(\nabla^2 \phi - div\Big(\frac{\nabla \phi}{\mid \nabla \phi \mid}\Big)\Big), \tag{10}$$

where δ is the derivative of H_ϵ. The local intensity force F_1 and the global intensity force F_2 are defined as:

$$F_1 = (1-\omega)[\int K_\sigma(y-x) \mid I(x) - f_1(y) \mid^2 dy - \int K_\sigma(y-x) \mid I(x) - f_2(y) \mid^2 dy], \tag{11}$$

$$F_2 = \omega[(I(x) - c_1)^2 - (I(x) - c_2)^2]. \tag{12}$$

In [15], the energy function of the local and global model in Eq. 10 is not convex, and the gradient descent solver can easily be trapped into a local minimum. Here we employ the global convex segmentation model [18] to make the energy function convex. Considering the gradient flow equation in the local and global Gaussian distribution fitting model in Eq. 10, we drop the last term and take $v = 1$, which yields:

$$\frac{\partial \phi}{\partial t} = \delta_\varepsilon(\phi)\Big(-(F_1 + F_2) + div\Big(\frac{\nabla(\phi)}{\mid \nabla(\phi) \mid}\Big)\Big). \tag{13}$$

Following the idea proposed by Chan et al. [18], the stationary solution of the above equation coincides with the stationary solution of:

$$\frac{\partial \phi}{\partial t} = -(F_1 + F_2) + div\Big(\frac{\nabla(\phi)}{\mid \nabla(\phi) \mid}\Big) = -(\lambda_1 e_1 - \lambda_2 e_2) + div\Big(\frac{\nabla(\phi)}{\mid \nabla(\phi) \mid}\Big). \tag{14}$$

where $e_i(x) = (1-\omega) \int K_\sigma(y-x) \mid I(x) - f_i(y) \mid^2 dy + \omega \mid I(x) - c_i \mid^2, \;\; i = 1, 2.$ Bresson et al. [19] transformed the constrained optimization problem to an unconstrained optimization problem by restricting the solution to lie within

a finite interval. This simplified flow represents the gradient descent for the minimization problem:

$$min_{a_0 \leq \phi \leq b_0} E(\phi) = min_{a_0 \leq \phi \leq b_0} \mid \nabla \phi \mid_1 + < \phi, r >, \tag{15}$$

where $r = \lambda_1 e_1 - \lambda_2 e_2$, and the restriction $a_0 \leq \phi \leq b_0$ is to guarantee a unique global minimum. Goldstein et al. [20] used the Split Bregman algorithm to solve the global convex minimization problem. In this work, we constrain ϕ as $0 \leq \phi \leq 1$, which can guarantee a unique global minima. The global convex model can be written as $min_{0 \leq \phi \leq 1} E(\phi) = min_{0 \leq \phi \leq 1}(\mid \nabla(\phi) \mid_1 + \langle \phi \cdot r \rangle)$.

Replace the standard total variation (TV) norm $TV(\phi) = \int \mid \nabla(\phi) \mid = \mid \nabla(\phi) \mid_1$ with the weighted version $TV_g(\phi) = \int g(\mid \nabla(\phi) \mid) = \mid \nabla(\phi) \mid_g$, where

$$g(\Theta) = \frac{1}{1 + \beta \mid \Theta \mid^2} \tag{16}$$

is the non-negative edge detector function, where β is a constant value.

After these steps, the minimization problem becomes:

$$min_{0 \leq \phi \leq 1} E(\phi) = min_{0 \leq \phi \leq 1} \mid \nabla(\phi) \mid_g + < \phi, r > . \tag{17}$$

The Split Bregman method is a fast method to solve the L1-Regularization problem. This approach has been widely used in the domain of signal restoration and image processing [21, 22, 23]. The basic idea of this method is to transform the constrained optimization problem to an unconstrained optimization problem. The Split Bregman algorithm for the minimization of the energy function in Eq. 17 was proposed by Yang et al. [17]. To apply the Split Bregman approach, an auxillary variable $\vec{d} \leftarrow \nabla(\phi)$ is introduced.

By introducing a Lagrange multiplier λ_L, the Bregman iteration is applied to strictly enforce the constraint $\vec{d} = \nabla(\phi)$, we obtain the following iterative steps:

$$\begin{cases} (\phi^{k+1}, \vec{d}^{k+1}) = argmin_{0 \leq \phi \leq 1} \mid \vec{d} \mid_g + < \phi, r > + \frac{\lambda_L}{2} |\vec{d} - \nabla\phi - \vec{b}^k|^2, \\ \vec{b}^{k+1} = \vec{b}^k + \nabla\phi^{k+1} - \vec{d}^{k+1}. \end{cases}$$

For fixed $\vec{d}$, minimize w.r.t. ϕ:

$$\triangle\phi = \frac{r}{\lambda_L} + \nabla * (\vec{d} - \vec{b}), \qquad 0 < \phi < 1. \tag{18}$$

Using central discretization for the Laplace operator and backward difference for the divergence operator, the numerical scheme is:

$$\begin{cases} \alpha_{i,j} = d^x_{i-1,j} - d^x_{i,j} + d^y_{i,j-1} - d^y_{i,j} - (b^x_{i-1,j} - b^x_{i,j} - b^y_{i,j-1} - b^y_{i,j}), \\ \beta_{i,j} = \frac{1}{4}(\phi_{i-1,j} + \phi_{i+1,j} + \phi_{i,j-1} + \phi_{i,j+1} - \frac{r}{\lambda_L} + \alpha_{i,j}), \\ \phi_{i,j} = max\{min\{\beta_{i,j}, 1\}, 0\}. \end{cases}$$

For fixed ϕ, minimize w.r.t. $\vec{d}$:

$$\vec{d}^{k+1} = shrink_g\Big(\vec{d}^k + \nabla\phi^{k+1}, \frac{1}{\lambda_L}\Big) = shrink\Big(\vec{d}^k + \nabla\phi^{k+1}, \frac{g}{\lambda_L}\Big), \tag{19}$$

where

$$shrink(x, r) = \frac{x}{\mid x \mid} max(\mid x \mid -r, 0). \tag{20}$$

The iteration termination criterion is defined as $\parallel \phi^{k+1} - \phi^k \parallel > \Psi$. Using this method, once the optimal ϕ is found, we can generate the segmented region $\Omega^k = \big\{x : \phi^k(x) > 0.5\big\}$. The implementation of the Split Bregman algorithm for minimizing the energy function in Eq. 17 can be summarized as follows:

1: while $\parallel \phi^{k+1} - \phi^k \parallel > \Psi$ do
2: Define $r^k = \lambda_1 e^k_1 - \lambda_2 e^k_2$
3: $\phi^{k+1} = GS(r^k, \overrightarrow{d^k}, \overrightarrow{b^k}, \lambda_L)$
4: $\overrightarrow{d}^{k+1} = shrink_g(\overrightarrow{b^k} + \nabla\phi^{k+1}, 1/\lambda_L)$
5: $\overrightarrow{b}^{k+1} = \overrightarrow{b}^k + \nabla\phi^{k+1} - \overrightarrow{d}^{k+1}$
6: Find $\Omega^k = \big\{x : \phi^k(x) > 0.5\big\}$
7: Update e^k_1 and e^k_2
8: end,

where $GS(r^k, \overrightarrow{d^k}, \overrightarrow{b^k}, \lambda_L)$ denotes the Gauss-Seidal iteration method, $\overrightarrow{b}$, $\overrightarrow{d}$ are auxiliary variables, $shrink_g$ is a shrinkage frame (see [19, 24, 25]), and Ω is the tumour region. We use this improved region-based active surface modelling method to segment VOIs in the PET volume.

2.3. Hierarchical Segmentation Scheme

Given the limited spatial resolution of PET images, we use the above segmentation in a hierarchical scheme that can effectively segment the VOI from the surrounding tissue. In the first step, we calculate the standard deviation value in a small window for each pixel. For each image, if the

average of all the standard deviation values is above an empirical threshold ψ, this image will be considered as a significantly noisy image, and will not be used to estimate the adaptive threshold T in Eq. 7. Next, we extract a closed surface area by using a low threshold to initialize the surface. Then the surface is deformed by minimizing the energy function in Eq. 3. The threshold T in Eq. 7 is estimated by calculating the average of the maximum grey level values of some selected slices in the volume. These slices, which contain potential MS lesions, are obtained by propagating a single slice which are most likely to contain MS lesions (see Fig. 3). Specifically, we obtain the maximum grey level value for each slice and normalize all the values between 0 and 1. We then select the slice which contains the maximum value (i.e. 1) as the starting point. This slice will be used to compare with its three preceding slices (named as left neighboring window). If the difference between the slice maximum and the maximum in subsequent slices within the left neighboring window is below a threshold φ, we move to its closest preceding slice and compare the slice with its corresponding left window. Otherwise, the propagation across the selected slices will be stopped. The same process is also imposed on the slices following the starting point. The selected slices are propagated in this way. The threshold T in Eq. 7 can be estimated by calculating the average maximum grey level values of these selected slices.

For robust segmentation, a reliability metric is used to remove false positive VOIs. Specifically, the 6-connected neighborhood voxels are labeled in the results. If the tumour only exists in one or two slices, the corresponding 3D labelling will be considered as a false positive VOI and will be removed from the results.

In the next step, we try to find the non-detected regions that are not selected in the first step. We use the local grey level volumetric information to segment the non-detected VOIs. Specifically, morphology is used to dilate each 3D VOI generated in the previous step. Then a 3D maximum bounding box which can be used as the initial closed surface is generated for each dilated VOI. Subsequently, each VOI is segmented again using the local grey level voxels. The threshold T in this step is estimated by computing the average grey level values of the voxels in each bounding box. The advantage of the approach lies in its multi-level processing. It first considers the whole range of grey levels in the volume and as such avoids being trapped into local minima. Secondly, the local range of grey levels is utilized to refine the segmentation which can effectively avoid false negative segmentations.

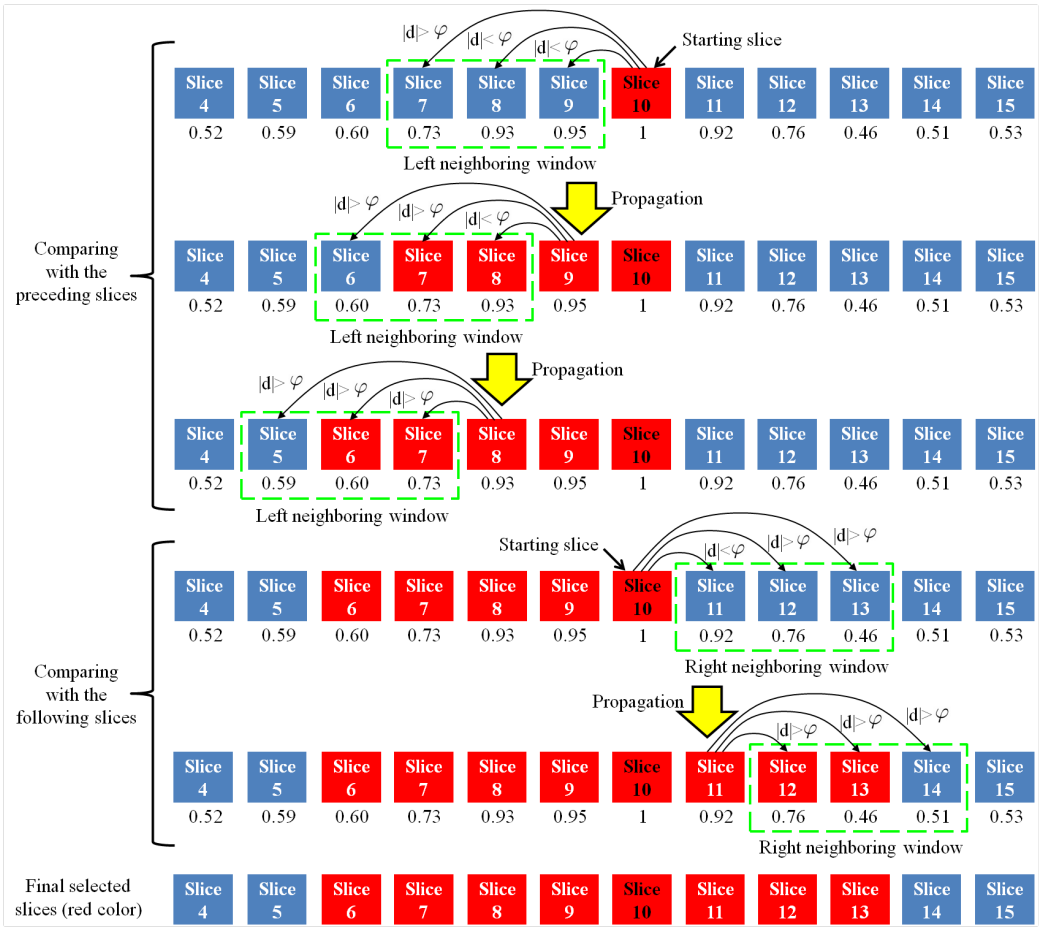

Figure 3: An example of the processing of the slice selection. The red color square indicates selected slices, the absolute value $|d|$ denotes the difference of the values between the two slices connected with arrow, the threshold φ is given as 0.1 in this example.

2.4. Segmentation Refinement by Alpha Matting

Due to the limited image resolution, active surface modelling may produce undesired or unrealistic deformed surfaces, especially for thin structures. In addition, the partial volume effect (when a pixel represents more than one kind of tissue type) blurs the intensity discontinuity between tissue classes at the border of abnormal tissue. The binary segmentation results generated from the previous steps might not accurately represent the real tumour boundaries. To further improve the segmentation accuracy, we introduce an alpha matting method [26] into our segmentation pipeline. We observe that

the blurry boundary of a VOI is caused by the fact that the boundary pixels are a mixture of the foreground tissue (tumours) and the background tissue (normal tissue). Instead of generating binary segmentation labels, alpha matting techniques can generate fractional alpha values between 0 and 1 for these pixels, which can be viewed as accurate soft segmentation.

To use alpha matting, a trimap has to be generated at first, which separates the image into three regions: definite foreground F, definite background B, and the unknown region U, as indicated in step 3 in Fig. 5 and Fig. 6. Our system automatically generates this trimap. Specifically, we use morphology to erode the previous segmentation result with a circular structuring element to obtain the foreground F (show in white), then the background B (show in black) can be generated by dilating the binary segmentation. The unknown area (show in grey) can be generated which contains all pixels that are not covered by either F or B. For some foreground regions which are very small, erosion will remove all the foreground pixels. To avoid this problem, we label the segmented regions in each slice and count the pixel number in each region. If the number of pixels in a labelled region is smaller than 9 (if the region is eroded by using a circular structure, it will disappear), the labelled region will be directly used as the foreground without erosion.

To solve the alpha matting problem, we use the approach proposed by Levin et al. [26]. Using this approach each single slice is modelled as $I_i = \alpha F_i + (1-\alpha)B_i$, where I is the observed image, F and B are foreground and background, and α is the transparency parameter. It can be rewritten as $\alpha_i = aI_i + b$, where $a = 1/(F-B)$, $b = -B/(F-B)$. Then the problem is converted to finding α, a and b to minimize the cost function:

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in w_j} (\alpha_i - a_j I_i - b_j)^2 + \varepsilon a_j^2 \right) \tag{21}$$

where w_j is a small window around pixel j. For $J(\alpha) = min_{a,b} J(\alpha, a, b)$, we have $J(\alpha) = \alpha^T L \alpha$, where

$$L_{i,j} = \sum_{k|(i,j) \in w_k} \left(\delta_{i,j} - \frac{1}{|w_k|} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\frac{\varepsilon}{|w_k|} + \sigma_k^2} \right) \right). \tag{22}$$

In this Laplacian function, $\delta_{i,j}$ is the Kronecker delta, μ_k and σ_k^2 are the mean and variance of the intensities in the window w_k around k which is usually 3×3, and $\mid w_k \mid$ is the number of pixels in this window. The details of the energy minimization process can be found in [26]. Solving the matting problem leads to a soft segmentation of VOIs in PET images.

3. Experiments

In this section, we present the test data and the experiments used to test the accuracy of the proposed hierarchical segmentation algorithm. We compare our approach with two state-of-the-art methods, namely PGVF [9] and MAP-MRF-EM [10].

3.1. Test Images

To evaluate the accuracy of the proposed method, we use 2 PET volumes of head-and-neck cancer patients and 2 volumes of a custom-built tumour phantom, as used in previous experiments at the XII Turku PET Symposium [27]. All images were acquired using a hybrid PET/CT scanner (GE Discovery) and the metabolic tracer [^{18}F]FDG. Volumes of interest are detailed in Tab.1. Each patient volume contains a tumour and in one case a metastatic lymph node. The phantom contains two glass VOIs representing tumour and lymph node, and was scanned twice with different activity FDG concentrations to reproduce typical signal-to-background ratios as seen in clinical practice.

Table 1: Details of two head and neck cancer PET images, and details of the phantom scans, with different activity concentrations in the VOI and background (BG).

PET image	VOI	PET protocol		
		initial activity (MBq ml^{-1})	frame (minutes)	voxel size (mm)
patient 1	tumour & node	240	3.0	2.73×2.73 ×3.27
patient 2	tumour	363	3.0	2.73×2.73 ×3.27
phantom scan 1	tumour & node	8.7 and 4.9 (VOI and BG)	10.0	1.17×1.17 ×3.27
phantom scan 2	tumour	10.7 and 2.7 (VOI and BG)	10.0	1.17×1.17 ×3.27

The patient images have the advantage of presenting a real segmentation problem, while the phantom images have the advantage of more reliable ground truth. In the case of phantom VOIs, ground truth masks were created from CT images that were acquired simultaneously by the PET/CT device and reconstructed to the same voxel grid as the PET images. The

CT images were thresholded at a density value (HU) corresponding to the transition between glass and tracer at the internal surface of the VOIs. The selected CT threshold recovered the known internal volume of the glass VOIs, measured independently from their liquid contents. In the case of patient tumours, ground truth segmentations were generated using manual delineation by three expert physicians based at the imaging site. Each expert delineated each of the three patient tumours twice.

In [28], Zeng et al. randomly selected one as the ground truth from the resulting 6 ground truth estimates. However, a single manual segmentation should not be considered as the ground truth as manual delineation is prone to variability [2, 29]. In this work, we generate a union ground truth by combining all the manual segmentation results together as:

$$UGD = \frac{\sum_1^N S_i}{N} \times 255, \tag{23}$$

where UGD denotes the union ground truth, S_i denotes one manual segmentation, and N is the total number of manual segmentations. Fig. 4 shows an example of the computed union ground truth, where Fig. 4(b-g) are the manually delineated segmentation results, and Fig. 4(h) is the union ground truth generated by combining all the manual results.

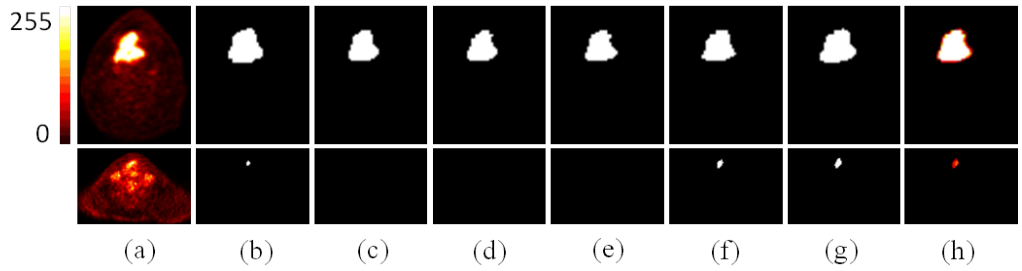

Figure 4: Examples of union ground truth for patient data. (a) patient PET data (slice 30 for patient 1, and slice 10 for patient 2), (b) manual result (physician 1 run 1), (c) manual result (physician 1 run 2), (d) manual result (physician 2 run 1), (e) manual result (physician 2 run 2), (f) manual result (physician 3 run 1), (g) manual result (physician 3 run 2), (h) union ground truth (the color bar is shown in the left corner).

3.2. Segmentation Results

We first describe the detailed segmentation process for the examples of a patient tumour and a phantom tumour which are shown in Fig. 5 and Fig. 6,

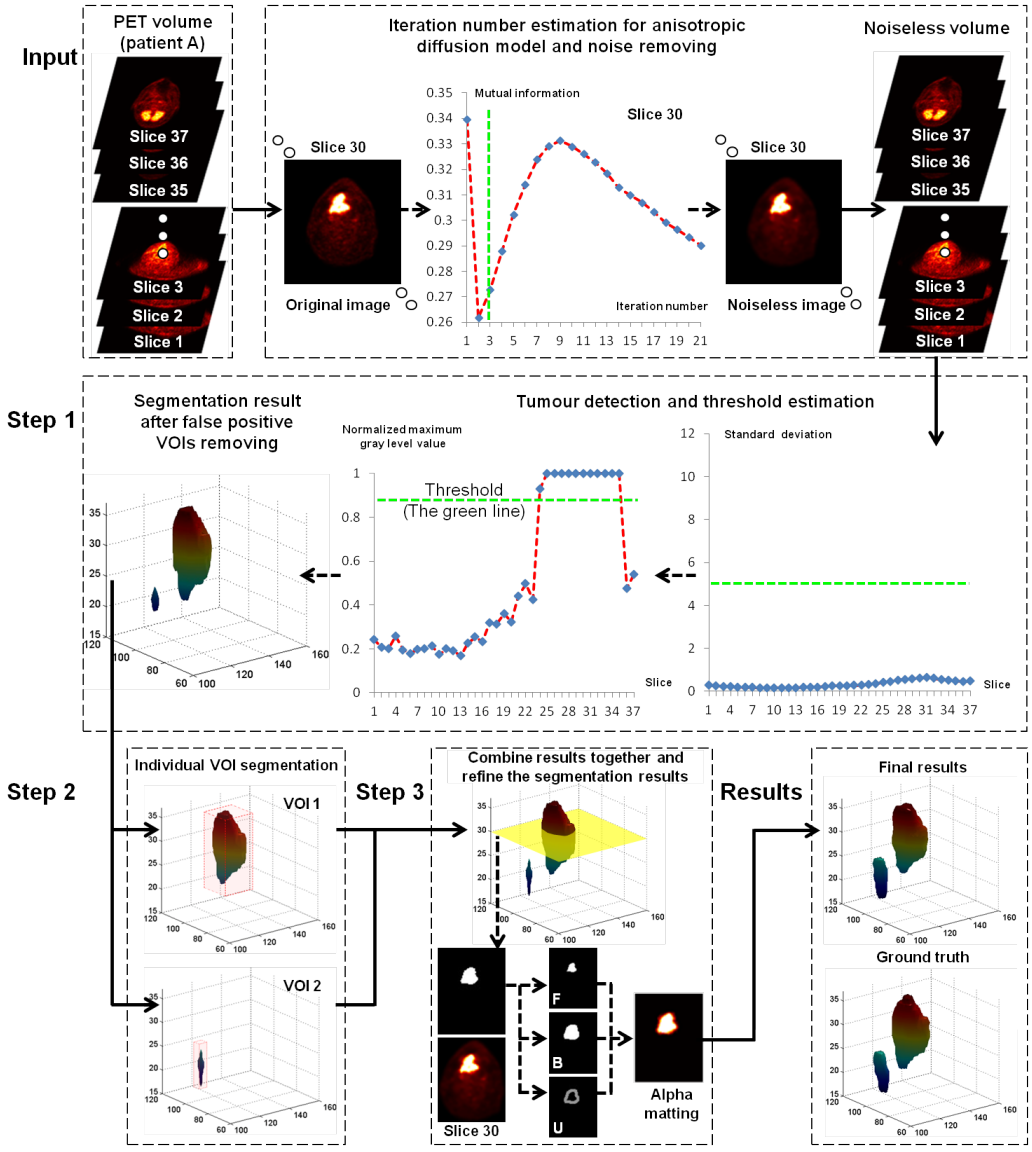

Figure 5: A patient example of tumour segmentation based on the proposed method.

respectively. In preprocessing, the anisotropic diffusion model is used to automatically reduce the image noise slice by slice. The iterative anisotropic diffusion will stop once the difference of the mutual information value is below τ, which is set to be 0.05. In this experiment, the automatically determined iteration number is 3 for patient slice 30 and 4 for phantom slice 20. In

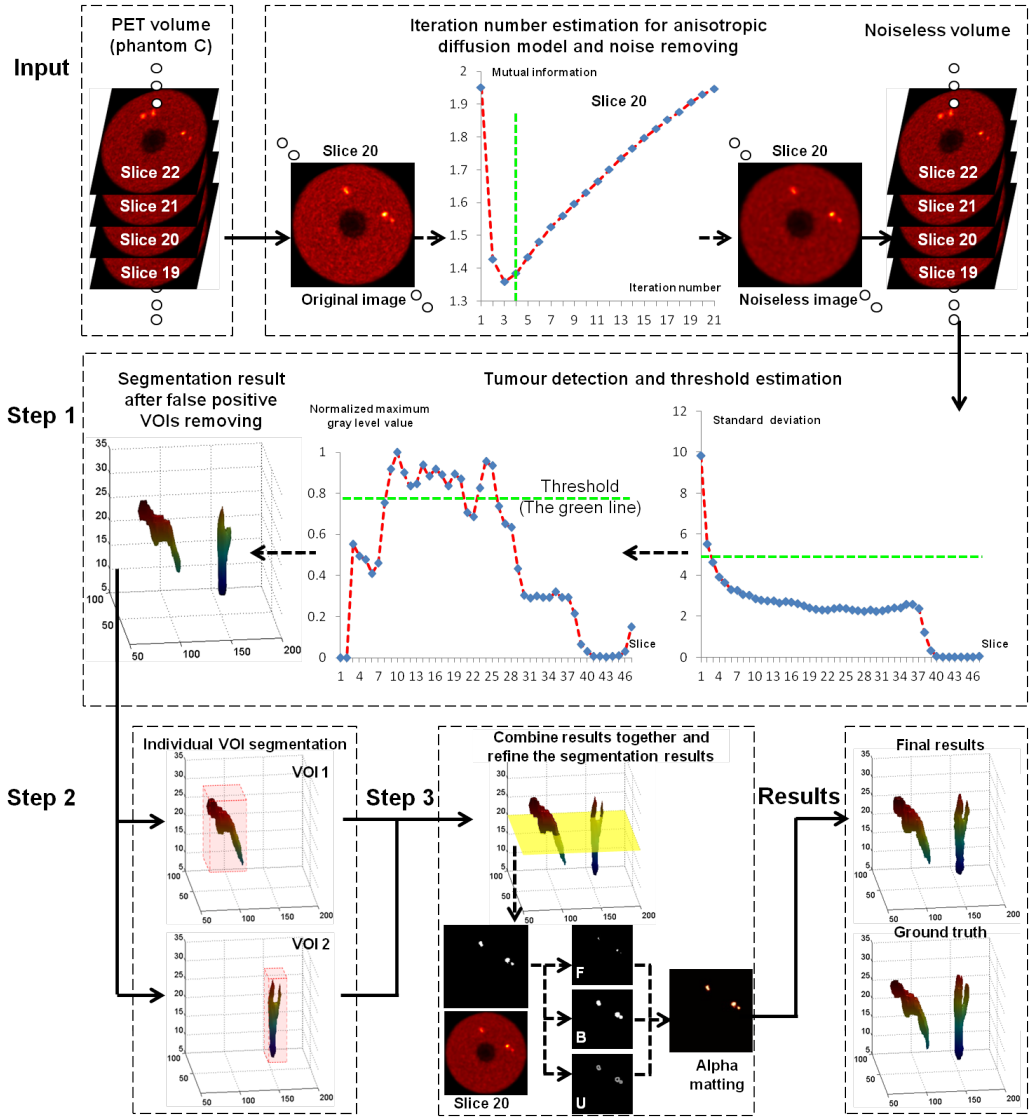

Figure 6: A phantom example of tumour segmentation based on the proposed method.

the first step, the standard deviation value is calculated in a 7×7 window for each pixel. We then calculate the average standard deviation for each slice in this volume. The slices which are below the threshold $\psi = 5$ will be selected to estimate the adaptive threshold T in Eq. 7. For the patient volume, the average standard deviation of all the slices are far less than ψ, so all the slices are used to calculate T. For the phantom volume, the first

and the second slices are above ψ, which significantly influences estimating the threshold T, so the first two slices will not be used to estimate T. We obtain all the maximum grey level values for each slice in the volume, and normalize all of them to be between 0 and 1. Subsequently, the selected slices (patient 1: from 23 to 36, phantom 1: from 7 to 30) are identified using the proposed criteria ($\varphi = 0.12$), then the threshold T in Eq. 7 is automatically estimated as 0.916 for the patient volume and 0.783 for the phantom volume, respectively. In this work, $\omega = 0.1$ in Eq. 4, the 3D Gaussian kernel σ is 3, $\lambda_i = 1$ in Eq. 5 and Eq. 6, $\xi = 0.1$ in Eq. 7, and $\lambda_L = 10$ in Eq. 18. The improved active surface modelling result is used to segment the image and false positive regions are automatically removed. In the second step, the previous results are dilated by using a sphere with radius equals to 10. The maximum bounding box is then generated as the initial surface for each VOI. Subsequently, the improved active surface modelling is used again to segment each bounding box, and the results are shown as semi-transparent red regions in step 2 in Fig. 5 and Fig. 6. The threshold T is recalculated for each dilated VOI. From the results we can see that some false negative regions are identified in this step. For example, the VOI2 in Fig. 5 step 2 and Fig. 6 step 2, respectively. In the third step, the alpha matting algorithm is used to further refine the segmentation. The radius for dilation and erosion for the trimap generation is fixed at 3 for all the images. Visual comparison with the ground truth, indicates the proposed method can accurately segment the VOIs.

We have compared our method with alternative approaches on real and phantom PET data. For real PET data, some images are randomly selected and are shown in Fig. 7. Compared with the ground truth, the results of PGVF and MAP-MRF EM methods have more false positives than our results, especially on slice 18 for patient 1. This is partially due to the low spatial resolution of the image, partial volume effects and the fact that these per-slice methods ignore volumetric grey level information. For the phantom data shown in Fig. 8, some of the images contain significant noise and density inhomogenity. The PGVF method fails to segment the abnormal region, since the density of the abnormal region is very similar to the surrounding tissue. The MAP-MRF EM method also does not perform well on images with significant noise. In contrast, our proposed method performs well on these images and provides accurate segmentation results.

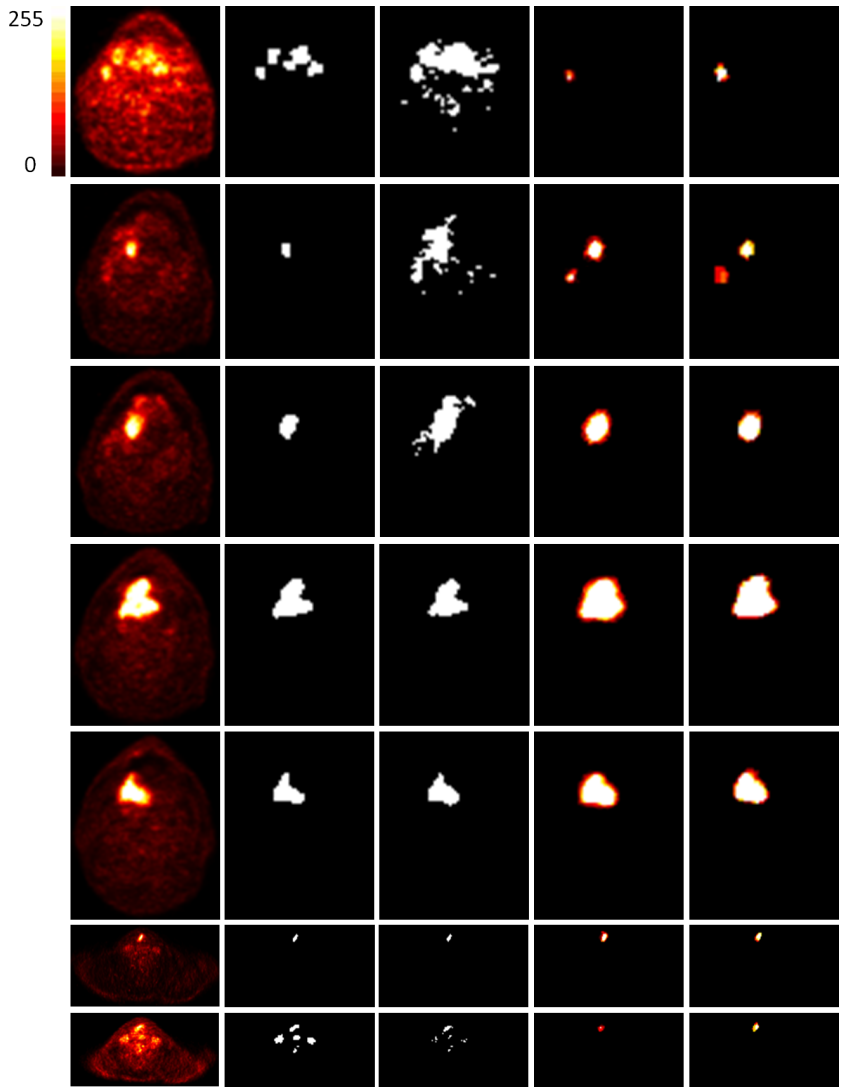

Figure 7: Segmentation of patient VOIs. From left to right: patient data (color bar in the upper left corner) (from top to bottom: slice 18, slice 24, slice 26, slice 30, slice 34 for patient 1, and slice 7, slice 10 for patient 2); PGVF [9]; MAP-MRF EM [10]; segmentation results of our method (color bar in the upper left corner); ground truth (color bar in the upper left corner).

3.3. Accuracy Evaluation

Previous work on comparing contouring methods in PET oncology [30, 31, 32] only use binary segmentation results as ground truth, thus cannot be

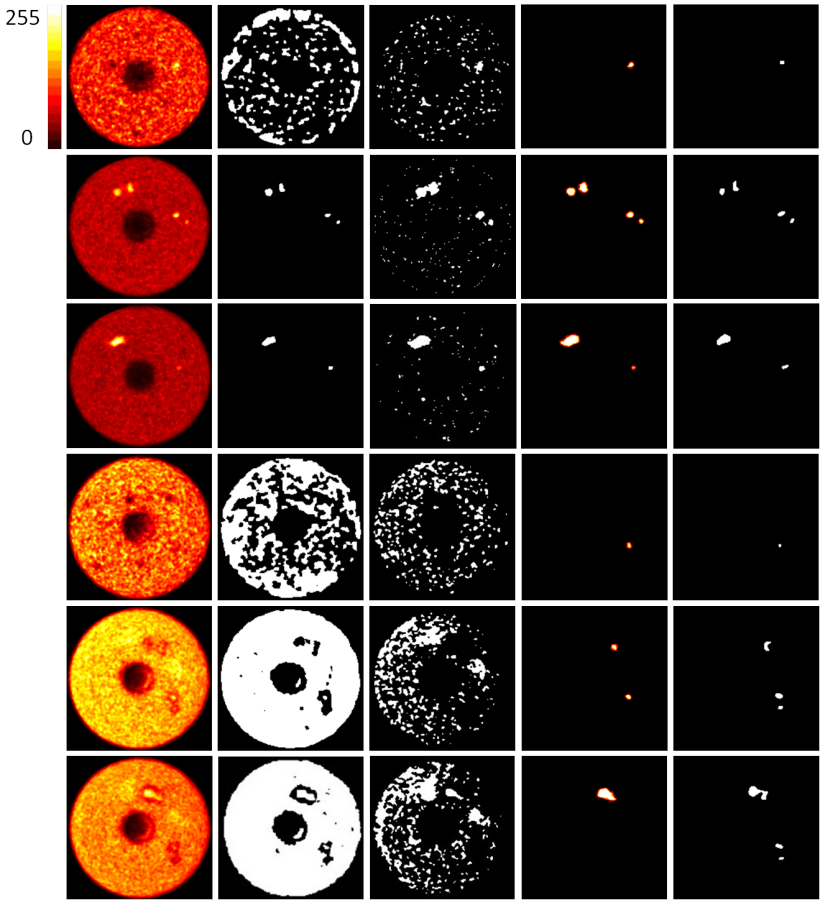

Figure 8: Segmentation of phantom VOIs. From left to right: phantom data (color bar in the upper left corner) (from top to bottom: slice 7, slice 22, slice 26 for phantom 1, slice 5, slice 19, slice 21 for phantom 2); PGVF [9]; MAP-MRF EM [10]; segmentation results of our method (color bar in the upper left corner); ground truth (color bar in the upper left corner).

used with the union ground truth. In this work, mutual information is used as a measure of statistical correlation of a segmentation result and the union ground truth. Given an image A (union ground truth) and B (segmentation result), their mutual information is defined as:

$$MI(A, B) = H(A) + H(B) - H(A, B), \tag{24}$$

where $MI(A, B)$ is the mutual information of image A and B, $H(A)$ and $H(B)$ are the entropy of image A and B, respectively, and $H(A, B)$ is the

Table 2: The mutual information distance and standard deviation ($MID \pm SD$) for tumour segmentation on PET data (all values $\times 10^{-3}$).

Algorithm	PGVF [9]	MAP-MRF EM [10]	Our method
Patient 1	3.10±3.76	3.32±4.34	1.33±1.42
Patient 2	5.50±1.35	4.43±1.11	4.18±1.04
Phantom 1	3.43±4.15	4.22±4.89	1.67±1.87
Phantom 2	5.89±7.28	5.60±6.80	3.98±5.14
All images	4.48±4.14	4.39±4.29	2.79±2.37

joint entropy of A and B. Fig. 9 shows the mutual information curves. The blue line denotes the mutual information of the ground truth $MI(A, A)$. For the patient images, the union ground truth is directly used to compare our method with alternative approaches [9, 10]. The mutual information curves of our method in patient images are closer to the ground truth curves, except for a few points for patient 2. For the phantom images, the mutual information curves of our method are always better than those of the other approaches.

For patient 2, it is difficult to evaluate the results using the mutual information curve, because some points of the proposed method are under the mutual information curve generated by the alternative methods, and the average of the mutual information value of our method is similar to the others. In this work, a mutual information distance (MID) for the VOIs between the ground truth $MI(A, A)$ and the segmentation results $MI(A, B)$ is proposed as:

$$MID = \frac{\sum_{i=1}^{N}(MI(A, A)_i - MI(A, B)_i)}{N}, \tag{25}$$

where N is the number of slices in the volume. The MID is used as a measurement for the distance between two curves (the curves of the segmentation result and the ground truth in Fig. 9). It is performed slice by slice rather than the whole volume in order to calculate the standard deviation for the VOIs in the whole volume. A lower MID value indicates a better segmentation result. In Table 2, we can see our MID values as well as the standard deviation (SD) are lower in all the cases (especially for patient 1, phantom 1 and phantom 2) as well as the average values (All images) compared to the other methods.

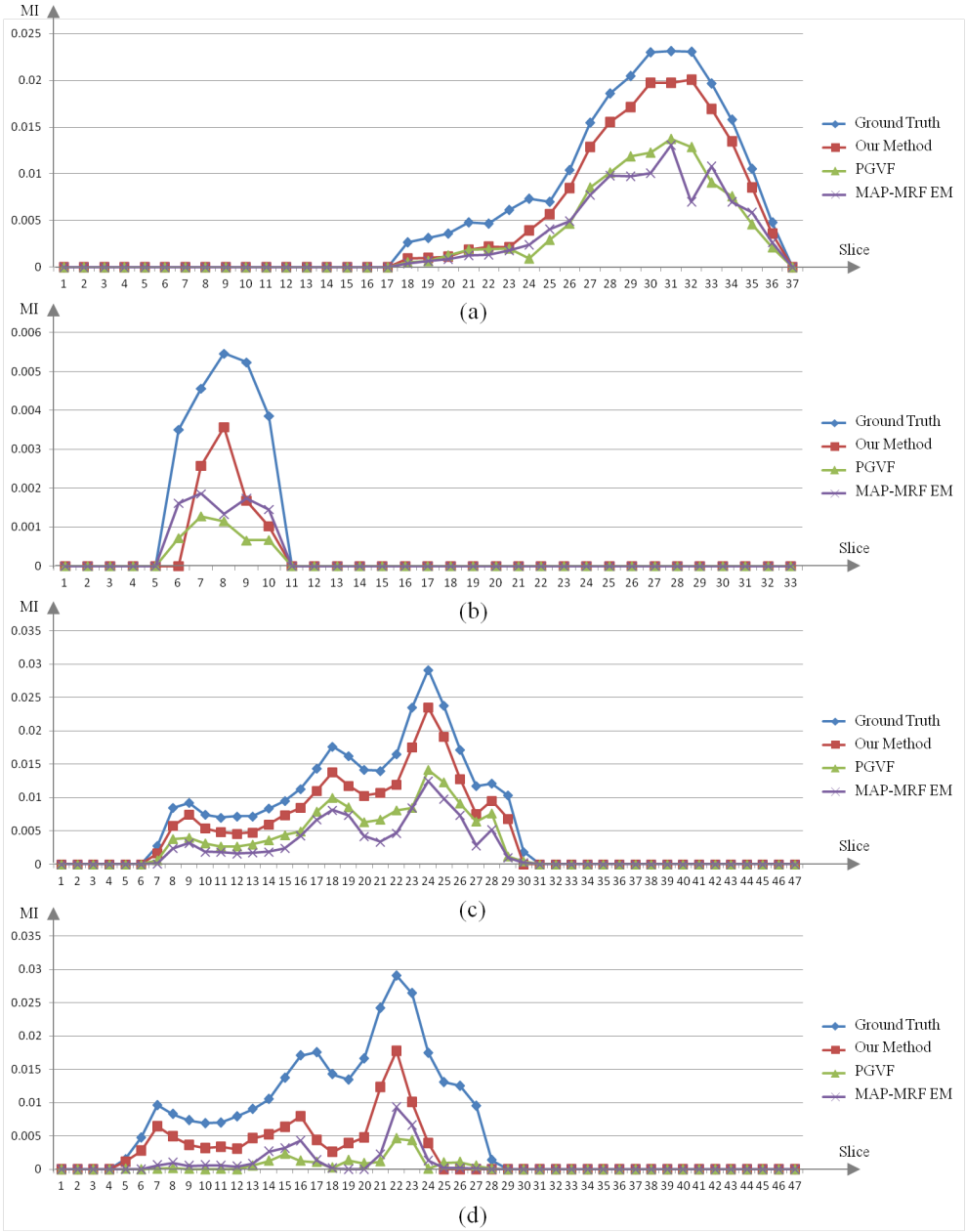

Figure 9: Mutual information for tumour segmentation on PET data. (a) patient 1, (b) patient 2, (c) phantom 1, (d) phantom 2.

To further evaluate the accuracy of the segmentation results, we measure the volumetric overlap of a segmentation with the ground truth. We use the standard overlap measure of the Dice similarity coefficient (DSC) [33] defined as:

$$DSC(A1, B1) = 2 \times (A1 \cap B1)/(A1 + B1), \tag{26}$$

where $A1$ and $B1$ denote the number of voxels contained in the ground truth and the segmentation result, respectively, and $A1 \cap B1$ denotes the number of voxels in the intersected region of $A1$ and $B1$. This measure has values in the range of 0 to 1, where DSC=1 indicates the exact overlap with the ground truth, i.e. the optimal segmentation. Using this metric we compare the proposed method and previous segmentation methods [9, 10] for both real and phantom PET. Note that computing DSC requires binary input images, but both our segmentation results and the union ground truth for the patient data have soft alpha values. We thus need to convert them from soft segmentation to binary segmentation for computing DSC values. To do so, we use a fixed threshold at the median grey level (128) for thresholding as the average of the union ground truth to calculate DSC for the patient data. Then we threshold the grey level of our results using different threshold values from 0 to 255 at intervals of 20. Fig. 10 shows the results of DSC and the standard deviation (SD), we can see that the results in DSC are close to the comparison methods [9, 10] at low threshold values except the phantom 2. As the threshold value increasing, our results in DSC are gradually increased and better than the alternative methods. From the average results in Fig. 10(e) and Fig. 10(f), we can see the values of DSC achieve to the maximum around the median threshold value. In contrast with the DSC, the best results of standard deviation (SD) can be obtained at the low threshold values. Although the standard deviations are gradually growing with the increasing of the threshold value, the average of the standard deviations in both patient and phantom volumes are better than the alternative methods [9, 10].

We also use a receiver operating characteristic (ROC) to evaluate spatial accuracy of the segmentation results. We denote TP as true positive, TN as true negative, FN as false negative, and TN as true negative. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a grey level threshold at intervals of size equals to 20 from 0 to 255. The ROC curve is created by plotting the fraction of true positive values (TPR=TP/(TP+FN)) and the fraction of false positive values (FPR=FP/(FP+TN)) at different grey level threshold. Accuracy is measured

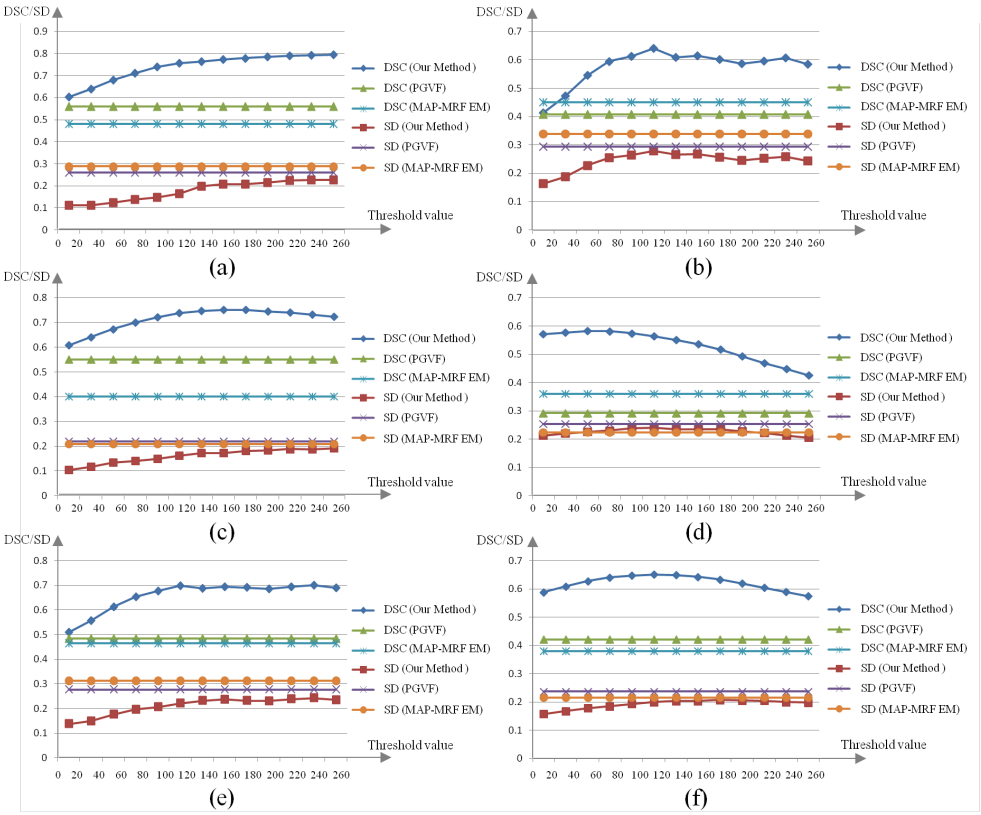

Figure 10: Dice index curves for tumour segmentation on PET data. (a) patient 1, (b) patient 2, (c) phantom 1, (d) phantom 2, (e) overall mean of patient PET, (f) overall mean of phantom PET.

by the area under the ROC curve. The best possible prediction method would yield a point in the upper left corner (0,1) of the ROC space, representing 100% accuracy. For the patient data, the ground truth is again thresholded at the median grey level (128). Fig. 11 (a, c, e, g) shows the ROC curves of all the test cases and Fig. 11 (b, d, f, h) show some zoomed-in parts of the ROC curves. We can see that the ROC curves of our method are steep due to low FP values. As the grey level threshold increases, our method shows a larger area under the ROC curve than the alternative approaches [9, 10]. The area under the ROC curve (AUC) is shown in Table 3, we can see the AUC values in our results are improved compared with [9, 10], and the standard deviation is better than [9].

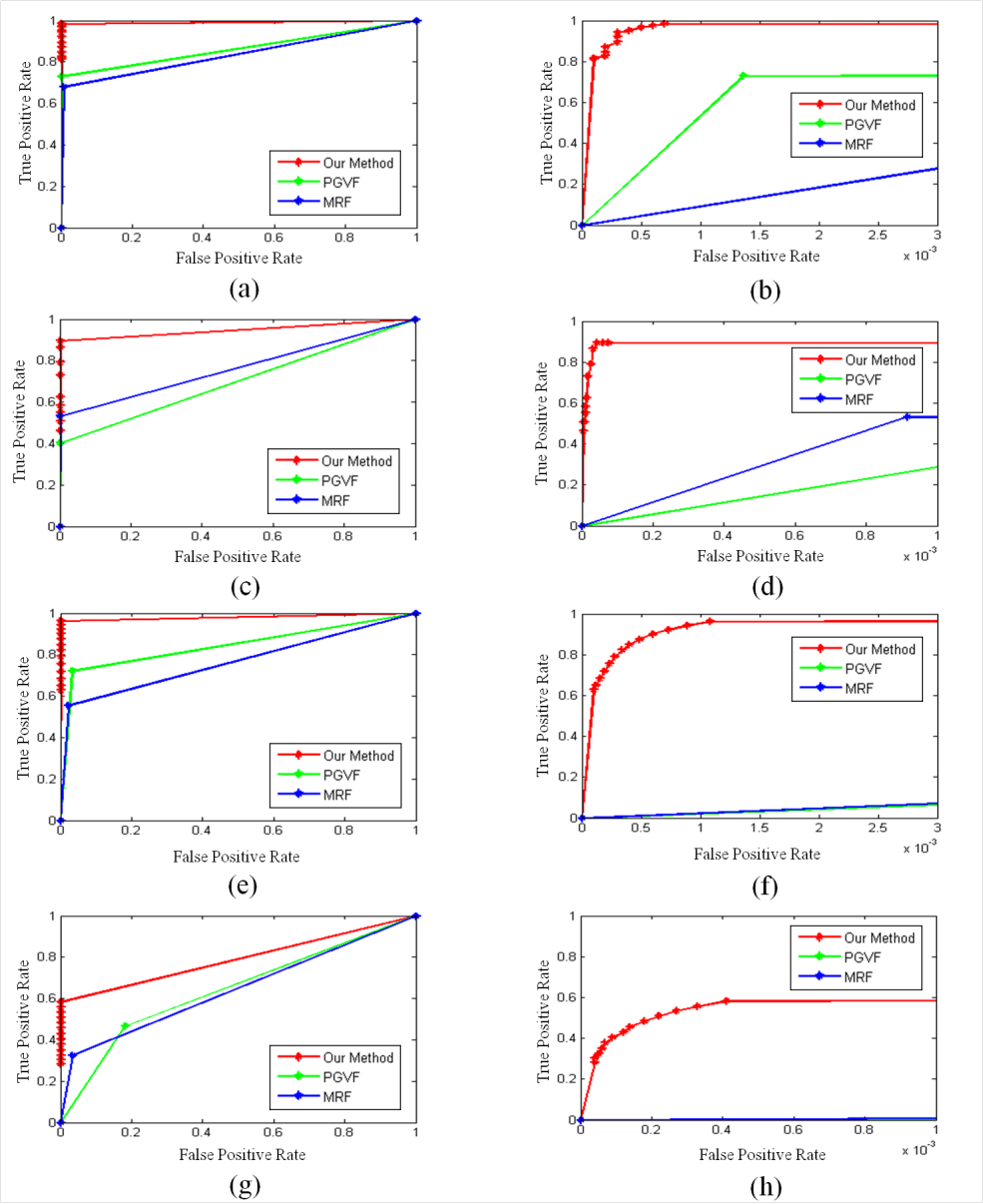

Figure 11: ROC curves for tumour segmentation on PET data. (a), (c), (e), and (g) are patient 1, patient 2, phantom 1, and phantom 2 ROC curve, respectively. (b), (d), (f), and (h) are zoomed-in parts of the patient 1, patient 2, phantom 1, and phantom 2 ROC curves, respectively.

Table 3: The area under the ROC curve (AUC) for tumour segmentation on PET data

Algorithm	PGVF [9]	MAP-MRF EM [10]	Our method
Patient 1	0.86	0.84	0.99
Patient 2	0.70	0.75	0.95
Phantom 1	0.85	0.77	0.98
Phantom 2	0.64	0.65	0.79
Average±SD	0.76±0.11	0.75±0.08	0.93±0.09

4. Discussion and Conclusions

In this paper, we introduce a new approach for PET image segmentation based on improved local and global intensity-based active surface modelling and alpha matting. The proposed method has the following four advantages. Firstly, it uses both local and global information to segment the volume data, leading to a more globally optimal solution and less segmentation errors. Secondly, a selection criteria for non-detected regions is used, which can effectively eliminate false negatives. Thirdly, the alpha matting technique is introduced to deal with partial volume effects, which achieves sub-pixel precision segmentation. Finally, a union ground truth for the patient data is introduced to evaluate our method using different measuring metrics. Evaluation shows that compared with alternative approaches, our method is more accurate and robust for PET image segmentation. As future work, we plan to extend the alpha matting method from 2D to 3D for improved handling of volumetric data. Also, we plan to further evaluate this method on a larger clinical PET dataset, and apply this method to other medical imaging problems.

References

[1] Y. Nakamoto, K.R. Zasadny, H. Minn and R.L. Wahl. Reproducibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[18F]Fluoro-D-Glucose. *Molecular Imaging Biology*, 4:171-178, 2002.

[2] S.L. Breen, J. Publicover, S. DeSilva, G. Pond, K. Brock, B. OSullivan, B. Cummings, L. Dawson, A. Keller, J. Kim, J. Ringash, E. Yu, A.

Hendler and J. Waldron. Intraobserver and interobserver variability in GTV delineation on FDG-PET-CT images of head and neck cancers. *International Journal of Radiation Oncology Biology Physics*, 68:763-770, 2007.

[3] Y.E. Erdi, O. Mawlawi, S.M. Larson, M. Imbriaco, H. Yeung, R. Finn and J.L. Humm. Segmentation of lung lesion volume by adaptive positron mission tomography image thresholding. *Cancer*, 80:2505-2509, 1997.

[4] B. Yaremko, T. Riauka, D. Robinson, B. Murray, A. Alexander, A. McEwan and W. Roa. Thresholding in PET images of static and moving targets. *Physics in Medicine and Biology*, 50:5969-5982, 2005.

[5] J.A.V. Dalen, A.L. Hoffmann, V. Dicken, and W.V. Vogel, B. Wiering, T.J. Ruers, N. Karssemeijer and W.J. Oyen. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. *Nuclear Medicine Communications*, 28(6):485-493, 2007.

[6] J.F. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux and V. Grégoire. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: infuence of reconstruction algorithms. *Radiotherapy and Oncology*, 69:247-250, 2003.

[7] U. Nestle, A. Schaefer-Schuler, S. Kremp, A. Groeschel, D. Hellwig, C. Rübe and C.M. Kirsch. Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, 34(4):453-462, 2007.

[8] S. Belhassen and H. Zaidi. A novel fuzzy c-means algorithm for unsupervised heterogeneous tumor quantification in PET. *Medical Physics*, 37:1309-1324, 2010.

[9] C.Y. Hsu, C. Liu and C. Chen. Automatic segmentation of liver PET images. *Computerized Medical Imaging and Graphics*, 32:601-610, 2008.

[10] H. Gribben, P. Miller, H. Wang, K. Carson, A. Hounsell and A. Zatari. Automated MAP-MRF EM labelling for volume determination in PET. In the proceeding of the IEEE International Symposium on Biomedical Imaging, 1-4, 2008.

[11] D.W.G. Montgomery, A. Amira and H. Zaidi. Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. *Medical Physics*, 34(2):722-736, 2007.

[12] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):529-539, 1990.

[13] C. Tong, S. Wang and L. Man. Improved image denoising method based on PDE. *Journal of Computer Engineering and Applications*, 46(15):176-179, 2010.

[14] Z. Wang and B.C. Vemuri. Tensor field segmentation using region based active contour model. *Lecture Notes in Computer Science*, 3024:304-315, 2004.

[15] L. Wang, C. Li, Q. Sun, D. Xia and C. Kao. Brain MR image segmentation using local and global intensity fitting active contours/surfaces. *Lecture Notes in Computer Science*, 5241:384-392, 2008.

[16] C. Li, C. Kao, C. John and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE Transactions on Image Processing*, 17(10):1940-1949, 2008.

[17] Y. Yang, C. Li, C. Kao and S. Osher. Split Bregman method for minimization of region-scalable fitting energy for image segmentation. *Lecture Notes in Computer Science*, 6454:117-128, 2010.

[18] T.F. Chan, S. Esedoglu and M. Nikolova. Algorithms for finding global minimizers of denoising and segmentation models. *SIAM Journal on Applied Mathematics*, 66:1632-1648, 2006.

[19] X. Bresson, S. Esedoglu, P. Vandergheynst, J. Thiran and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28:151-167, 2007.

[20] T. Goldstein, X. Bresson and S. Osher. Geometric applications of the Split Bregman method: segmentation and surface reconstruction. *Journal of Scientific Computing*, 45:272-293, 2010.

[21] T. Goldstein and S. Osher. The Split Bregman method for L1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323-343, 2009.

[22] X. Tai and C. Wu. Augmented Lagrangian method, dual methods and split bregman iteration for ROF model. *Lecture Notes in Computer Science*, 5567:502-513, 2009.

[23] S. Setzer, G. Steidl and T. Teuber. Deblurring Poissonian images by split Bregman techniques. *Journal of Visual Communication and Image Representation*, 21:193-199, 2010.

[24] C. Li, C. Kao, J. Gore and Z. Ding. Implicit active contours driven by local binary fitting energy. In the proceeding of Computer Vision and Pattern Recognition, 1-7, 2007.

[25] S. Osher and J.A. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12-49, 1988.

[26] A. Levin, D. Lischinski and Y. Weiss, A closed form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):228-242, 2008.

[27] T. Shepherd, M. Teräs, R. Beichel, R. Boellaard, M. Bruynooghe, V. Dicken, M. Gooding, P. Julyan, J. Lee, S. Lefèvre, V. Naranjo, U. Nestle, X. Wu, H. Zaidi, Z. Zeng and H. Minn. Comparative study with new accuracy metrics for target volume contouring in PET image guided radiation therapy. *IEEE Transactions on Medical Imaging*, 32(11):2006-2024, 2012.

[28] Z. Zeng, T. Shepherd and R. Zwiggelaar. Hierarchical modelling for unsupervised tumour segmentation in PET. In the proceeding of the IEEE-EMBS International Conference on Biomedical and Health Informatics, 439-443, 2012.

[29] A.C. Riegel, A.M. Berson, S. Destian, T. Ng, L.B. Tena, R.J. Mitnick and P.S. Wong. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *International Journal of Radiation Oncology Biology Physics*, 65:726-732, 2006.

[30] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rübe and C. Kirsch. Comparison of different methods for delineation of 18F-FDG PET positive tissue for target volume definition in radiotherapy of patients with nonsmall cell lung cancer. *Journal of Nuclear Medicine*, 46(8):1342-1348, 2005.

[31] C. Greco, S.A. Nehmeh, H. Schöder, M. Gönen, B. Raphael, H.E. Stambuk, J.L. Humm, S.M. Larson and N.Y. Lee. Evaluation of different methods of 18F-FDG PET target volume delineation in the radiotherapy of head and neck cancer. *American Journal of Clinical Oncology*, 31(5):439-445, 2008.

[32] H. Vees, S. Senthamizhchelvan, R. Miralbell, D.C. Weber, O. Ratib and H. Zaidi. Assessment of various strategies for 18F-FET PET-guided delineation of target volumes in high-grade glioma patients. *European Journal of Nuclear Medicine and Molecular Imaging*, 36:182-193, 2009.

[33] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26:297-302, 1945.